

CHAPTER 15

HOW LEXICAL IS MORPHOLOGY? THE CONSTRUCTICON AND THE QUADRIPARTITE ARCHITECTURE OF GRAMMAR

LIVIO GAETA

Abstract

Following Aronoff (1994), at least two different senses of the term “lexicon” must be distinguished. The Bloomfieldian sense of the term generally refers to the set containing any sort of entrenched or idiomatic expressions, while the second sense refers to the infinite “set of potential (regularly derived or compounded) lexemes for any given language.” A theory of lexeme formation makes crucial reference to this second sense and actually should keep it sharply distinct from the first one because it is only this latter that constitutes its real object of investigation. In this paper, this view will be taken seriously as a vantage point from which the relation between the two senses of the lexicon will be investigated. It will be shown that apparent paradoxes given by reduced phrases, phrasal compounds and coordination reduction, far from representing negative evidence, obey a clear ratio which neatly emerges if the multi-faceted perspective of the Constructicon is adopted as the interface of the different modules of grammar.

Key words

Please insert 10 keywords separated by commas and with minimal capitalisation

1. Introduction

In current morphological theory it is not clear what role has actually to be attributed to the lexicon. While much effort has been spent for assessing basic issues such as the distinction between inflectional and derivational morphology or compounds and phrases, much less discussion has been devoted to the relation between morphology and the lexicon. Although in the 1970s a Lexicalist view was felt to be justified as a reaction against the syntactic approaches of the 1960s, it has never been really clarified how the lexicon has to be concretely conceived. This question firstly concerns the basic units which constitute the backbones of the lexical component and enter word-formation patterns. In this light, it is not clear whether the word intended as a lexeme or lexical unit actually represents a relevant unit for morphology. In this paper, I will try to shed some light on this and other issues with the aim of finding out principles for establishing a consistently lexical morphology. In section 2 the issue of the role of the lexicon for morphology will firstly be tackled by showing that two different views of the lexeme have to be distinguished. The relevance of these two meanings of the lexeme will be concretely shown in section 3, in which the usage-based idea of the Constructicon will be seriously assumed, while in section 4 two complex cases in which morphology and syntax are strictly intertwined will be shown to support the principle of the lexuality of the input. The final section 5 draws the conclusions discussing the idea of a quadripartite architecture of grammar in which morphology has a (semi-)autonomous status with regard to syntax.

2. Is morphology lexical(ist)?

Far from being trivial, the question whether and to what extent morphology has to be considered lexical has received radically opposite answers in the current theoretical debate. This is the case of the well known controversy between Transformationalists and Lexicalists which goes back to the early years of Generative Grammar in the 1960s (cf. Scalise and Guevara 2005 for an historical reconstruction). In spite of several decades in which the Lexicalist paradigm seemed to have become prevalent, there is in fact a considerable number of fierce opponents of the Lexicalist view who defend a syntactocentric view of grammar and are correspondingly inclined to see the levels of morphology and syntax as strictly intertwined. This is notably the line adopted within Distributed

Morphology, in which the conflict between syntax and the lexicon is clearly solved in favor of the former (Embick and Noyer 2007, 289):

[a] theory of the syntax/morphology interface is first, a theory of how ‘words’ and their internal structure—the traditional domain of *morphology*—relate to the structures generated by the syntax, and second, a theory of how the rules for deriving complex words relate to the rules for deriving syntactic structures.

The terrain where this conflict manifests itself in its most dramatic consequences is the concept of word. In fact, in Distributed Morphology words are assembled by rules of the syntax and therefore don’t constitute a privileged derivational object as far as the architecture of the grammar is concerned, since all complex objects, whether words and phrases, are treated as the output of the same generative system (the syntax). One advantage of this approach is in terms of economy of derivation since it allows for a transparent (or direct) interface between syntax and morphology. In the default case, then, the principles that govern the composition of “words” are the same of those that govern the composition of larger syntactic objects. One can add that this approach views the lexicon in the traditional, Bloomfieldian way as a rather impoverished store in which only those objects are contained which cannot be generated by the derivational component. These are abstract morphemes or vocabulary items which are not complex in any meaningful sense, e.g. roots such as *√destroy* which gives rise to derivatives such as *to destroy*, *destruction*, *destructive*, *destroyable*, etc., contextually implemented in the corresponding syntactic environments.

Such an impoverished view of the lexicon is meant to strongly contrast with the Lexicalist approach which maintains that (at least some) words are special in ways that e.g. phrases are not, and that this specialness calls for an architecture in which the derivation of words and the derivation of syntactic objects occur in different modules of the grammar (the lexicon versus the syntax). While the words derived in the lexicon serve as the terminals in the syntactic derivation, there is a sharp division between syntax and morphology according to Lexicalist approaches of this type.

The interface between syntax and morphology in such a theory is opaque or indirect: there is no reason to expect the structure and composition of words to relate to the structure and composition of syntactic objects in any transparent or systematic fashion. In fact, a number of mismatches is usually shown to justify the autonomy of morphology. One of the mostly cited examples are the linking elements occurring in Germanic compounds like German *Liebe-s-brief* ‘love letter’,

in which the linking element cannot be explained away as a contextually determined form (e.g. a genitive) of the inflectional paradigm of the noun *Liebe* ‘love’. On the other hand, an indirect correspondence of a diachronic nature between syntactic and morphological objects can be at least partially expected if we consider that often morphological structure results from the grammaticalization of earlier syntactic structure (cf. Gaeta 2004, 2008, Wischer 2011). The watchdog enrolled to control the independence of the word from syntactic operations is the Lexical Integrity Principle: “[...] the syntax neither manipulates nor has access to the internal form of words” (Anderson 1992, 88). This principle, declined as it is in this classical formulation, responds to one general goal: defining canonical wordhood on the basis of the criterion of cohesiveness or non-interruptability. As we will see in section 2.4 below, although it is often taken for granted, the criterion of cohesiveness turns out to be illusory. Much worse, this illusion paves the way to the severe criticism of those who completely discard the word as a relevant object of the language. But before going into these details, let us understand what is meant by “word” in a Lexicalist approach.

2.1 What is the role of the word?

Clearly, in the Lexicalist view words play an important role as they contrast with phrases insofar as the latter result from syntactic operations. As it often happens to be in any science, there are many ways to deal with basic concepts like this. One widespread way of moving away from such a slippery basic concept is to replace it with an abstract notion like, in our case, the lexeme. Then, the next following questions sound: what is the relation between words and lexemes? What are relevant for a Lexicalist view, words or lexemes? As is well known, a big debate was started off by Aronoff’s (1976) classical book in whose preface a clear preference for the word was declared (Aronoff 1976, xi, emphasis added):

I have avoided the term *lexeme* **for personal reasons** and use instead the term *word*. This means that I have no way of distinguishing an uninflected word (lexeme) from an inflected word (word). I am confident that the ambiguity will not cause much grief.

We are not told about the personal reasons, but in the subsequent decades different positions were defended in support or against the word-based view. For instance, Scalise (1984, 59) discusses the relevance of a morphology based on concrete words rather than on pieces of words and specifically morphemes as pled for by others. Assuming such a view

forces Scalise to adopt a number of strategies to deal with typical instances of stem-based derivations as for instance rules of truncation which are necessary to correctly derive Italian suffixations like *fama* ‘fame’ > *famoso* ‘famous’ instead of the expected **famaoso* found for instance in *virtù* ‘virtue’ > *virtuoso* ‘virtuous’. Actually, this machinery, which can appear imposing in a consistently stem-based language like Italian, turned out to be mostly useless because Aronoff regretted later to have put such a big emphasis on words and confessed to intend his approach as based on lexemes instead (Aronoff 1994, 7, emphasis added):

[i]n Aronoff 1976 (henceforth *WFGG*) I used the term *word* in several senses and specifically noted in the preface that I would not use the term *lexeme*. This refusal led to a number of problems. For instance, one of the major points of *WFGG* was that morphology was what I termed *word-based*, **by which I meant lexeme-based**.

“Lexeme” is used by Aronoff in the traditional way as an abstract unit, deprived of any inflection. Therefore something like <fam + a>, which renders the adoption of truncation rules at least redundant, since they have been basically introduced to account for the stem-based nature of derivational rules in Italian. This is made explicit not without a certain embarrassment as he also reveals the personal reasons causing his previous dismissal of the term “lexeme” (Aronoff 1994, 7, emphasis added):

I especially did not mean that the **base** or **stem** for a word formation rule had to be **a complete word or free form**, only that **the base should be a lexeme** ... Nevertheless, others naturally misunderstood my claim as being about the forms of stems and pointed out that there are many languages in which **the actual form** to which a morphological operation applies **is often not a free form**, which would thus falsify my apparent claim that a stem had to be an otherwise free form (word). This was understandable, but even more to the point was the fact that **this particular homonymy confused me too**.

Trying to get rid of this confusion, Aronoff apostatizes and adopts the lexeme as a primary unit of morphology whereas it is now the term “word” that causes confusion (Aronoff 1994, 14):

I do not use the more common term word formation, because, unlike lexeme formation, it is confusing. That is because the term word is ambiguous among quite separate and independently important concepts.

However, things are not doing as well as hoped since also the term “lexeme” needs to be used with a certain caution, as we will discuss in the next section.

2.2 What is really a lexeme?

A standard definition of a “lexeme” pertains to its lexical status: under “lexeme₁” we can understand a unit which is stabilized or entrenched in the lexicon. Thus, for instance, no one would contend that the nominalized infinitive *pouvoir* ‘power’ of the following French sentence is a lexeme by itself distinct from its verbal cognate:

- (1) *Son département économique s’est penché sur l’efficience de nos pouvoirs publics*
 ‘His economic department has studied the efficiency of our public powers.’

As shown by the example, the infinitive is actually lexicalized, i.e. it has become a unit of the lexicon: we can label this sense “lexeme₁”, is shown by its inflectional behavior which is assimilated in full to that of simple nouns. Accordingly, we distinguish between a *<pouvoir>_v* and a *<pouvoir>_N*. On the other hand, we also normally use the term “lexeme” for any unit resulting from a process of lexeme-formation: we can call this sense lexeme₂. The question explodes in all its dramatic consequences when we consider cases in which both senses of lexemes occur close to each other. For instance, in the following French sentences we can use two different kinds of nominalization:

- (2) *L’idéologie de la beauté/du beau_N a été toujours interprété(e) comme l’expression d’un bel_{Adj} esprit qui se manifeste dans toutes les occasions les plus significatives qu’on rencontre dans les belles_{Adj} lettres.*
 ‘The ideology of beauty/the beautiful has always been interpreted as the expression of a beautiful mind who manifests itself in all the most significant occasions that are found in the belles-lettres’.

While no one would contend the lexeme₂ status of *beauté* as this results from the suffixation of the base adjective *beau*, it is unclear what to do with the nominalization *beau_N* which parallels *poivoir_N* seen above. Is *beau_N* to be considered as a lexeme₁ or as a lexeme₂, respectively on a par with *poivoir_N* and *beauté*, or rather a word form of *beau_{Adj}*? Notice that

this question also has a direct bear on the number of occurrences of the lexeme *beau*_{ADJ} in the text, whether two or three. In the Lexicalist approach the traditional answer distinguishes a morphological from a syntactic nominalization:

- a) Morphological nominalization: $\langle beau \rangle_{ADJ} > \langle beauté \rangle_N$
- b) Syntactic nominalization: $\langle beau \rangle_{ADJ} > [beau]_N$

The latter does not give rise to lexemes₂ in contrast with the morphological nominalization. Accordingly, $[beau]_N$ is not a lexeme₂ because it does not result from a morphological process. Moreover, $[beau]_N$ is not a lexeme₁ either because in contrast with $\langle pouvoir \rangle_N$ it cannot be said to be stabilized in the lexicon.

Notice incidentally that Distributed Morphology apparently does not have this problem because both *beauté* and *beau*_N contextually implement an abstract root \sqrt{beau} , any distinction with regard to the lexical status being irrelevant in that model. On the other hand, some lexical information needs to be added in that model too, insofar as $\langle pouvoir \rangle_N$ has to be treated as an idiom whose meaning is not wholly predictable from its morphosyntactic structural description. The same is true however for *beauté* as used in the following example from the Internet (ref??):

- (3) *Le marché du lundi au bar Rimini–bain de rivière pour hommes le jour–est une bonne adresse pour acheter vêtements et accessoires. Ici, on peut se faire coiffer, y observer et y rencontrer des beautés jusqu’à minuit.*
 ‘The Monday market at the bar Rimini–river bath for men during the day–is a good address where to buy dresses and accessories. Here, one can get the hairs dressed but also observe and meet beauties until midnight’.

The meaning extension $|beauty| > |beautiful\ woman|$ is clearly contextually determined but at the same time idiosyncratically connected with that specific idiom as for instance it is not valid for *beau*_N. Thus, also Distributed Morphology is forced to attribute to the lexicon a much larger role than the simple store for abstract roots. At any rate, this example shows that in truth a lexeme₂ tends to become a lexeme₁. In fact, we find in the literature the apparently paradoxical expression “lexicalized lexeme” which is meant to exactly reflect this phenomenon (Bauer 1983, 48, emphasis added):

[t]he final stage comes when, because of some change in the language system, the lexeme has, or takes on, a form which it could not have if it had arisen by the application of productive rules. At this stage **the lexeme is lexicalized**.

In particular, Bauer refers to single cases which result from a completely unproductive morphological process and are accordingly listed in the lexicon (Bauer 1983, 93, emphasis added):

[c]ertain word-formation processes are triggered or limited by the individual roots [...] the ending *-ric* only occurs in conjunction with *bishop* in English. This is an extreme example, which would probably be more economically captured by listing BISHOPRIC as **a lexicalized lexeme**.

Clearly, one cannot be satisfied with such a contradictory terminology, which implies that lexicalized lexemes are opposed to non-lexicalized ones. In fact, also Bauer in a footnote admits to be not eager to adopt it, although it is perfectly coherent with the ambiguity intrinsically present in the concept of lexeme (Bauer 1983: 48, emphasis added):

The term is no doubt unfortunate because of its other technical meanings [...]. Nonetheless it seems to have currency in studies of word-formation in approximately the sense in which it is being used here).

At any rate, Aronoff is the first who really puts the finger on the wound and calls for a neat distinction of the two senses (Aronoff 1994, 10, emphasis added):

[a]ll vocabulary words that are members of a major lexical class, regardless of whether they are actual or potential, are lexemes. The set of potential (regularly derived or compounded) lexemes for any given language is therefore infinite [...] **Being a lexeme and being in a (Bloomfieldian) lexicon are thus separate matters.**

Accordingly, we need to distinguish the lexicon₁ (= Lex₁), which consists of a list of form/meaning arbitrary pairs coming from different procedures and sources as it contains all expressions presenting idiomatic traits, including those larger than one word, from the lexicon₂ (= Lex₂) consisting of the (infinite) set of all potential lexemes and therefore resulting exclusively from lexeme-formation procedures. Actually, Aronoff's distinction aims at completely dismissing Lex₂ in favor of Lex₁ (Aronoff 1994, 22, emphasis added):

[...] since the extensional notion of a potential lexicon plays no significant role that I know of in any theory of morphology, while the Bloomfieldian theory is crucial to a proper understanding of blocking and productivity, it seems best to simply dispense with both the notion of the potential lexicon

and the name. The term *lexicon* should therefore be understood [...] as referring only to the permanent lexicon: the list of all idiosyncratic signs, regardless of their category or complexity. **The endless list of all lexemes, by contrast, will remain nameless.**

Therefore, Aronoff radically dispenses with the two-faced character of the lexeme and rejects Bauer's lexicalized lexemes. The implicit assumption, however, is that any (also potential) product of word-formation is immediately part of the lexicon on a par with cases like *bishopric* or, as an alternative, dismissed in a nameless land. At any rate, it is far from being clear whether the Lex₂ does not play any role in the quite complex issues of blocking and productivity as will be discussed in the following section.

2.3 The active role of the Lex₂

Under lexical blocking at least two different phenomena are comprised. By making reference to the role played by the Lex₁ (the Bloomfieldian lexicon), Aronoff only considers the so-called word-blocking, termed "Paul-blocking" in Gaeta (2015a). This refers to the blocking effect of a Lex₁-lexeme on the potential result of a word-formation rule, namely a Lex₂-lexeme:

- (4) a. Paul-blocking: Lex₁ > Lex₂
 b. *thief* > ??*stealer*

The occurrence of the Lex₁-lexeme *thief* blocks the potential formation of the Lex₂-lexeme °*stealer*, which, far from being ill-formed, is in fact found in compounds like *scene-stealer*. By doing so, Aronoff completely disregards a second type of blocking, the so-called rule-blocking, termed "Pāṇini-blocking" in Gaeta (2015a):

- (5) a. Pāṇini-blocking: Lex₂ > Lex₂
 b. [...]_{ADJ} > [...]_{ADJ} -*heit*_N
ordinär > *Ordinärheit*/??*Ordinarität*
 'vulgar' > 'vulgarity'
 c. [+learned]_{ADJ} > [...]_{ADJ} -*ität*_N
binär > *Binarität*/??*Binärheit*
 'binary' > 'binarity'

In this case, it is the more specific selective context of a word-formation pattern that blocks the application of a less specific one, as in the case of

the German deadjectival noun *Binarität* blocking the formation of the Lex₂-lexeme ??*Binärheit*, since *-ität* usually selects adjectives belonging to the learned vocabulary such as *binär*. This is shown by the block of the formation of the Lex₂-lexeme ??*Ordinarität* as *ordinär* does not match the specific property required. Being the result of the interaction of word-formation patterns, these reciprocal blocking effects can only be understood at the level of the Lex₂, although the patterns can also make reference to information coming from the Lex₁, such as usage-based features like [\pm learned], [\pm frequent], etc. Accordingly, the constraint on learnedness can be overcome if an adjective turns out to increase in usage frequency becoming part of a more colloquial register, as shown for instance by *skurril* ‘droll’ > *Skurrilheit* ‘drollery’, beside established *Skurrilität* (cf. Gaeta 2015a, 867). On the other hand, *Ordinarität* is found as a special term of mathematics with the meaning: ‘the property of being a common event’, clearly derived from the base *ordinär* ‘common, with regard to events’.

Furthermore, the effects of the Lex₂ can also be observed in the other domain considered irrelevant by Aronoff, namely productivity. As is well known, this concept comprises very different dimensions and phenomena (cf. Bauer 2001, Gaeta and Ricca 2015 for surveys). One core aspect refers to its dynamics intended as the probability of application of a derivational pattern, which can be captured in quantitative terms by the following formula (cf. Baayen 2009):

$$P = h_{Af} / N_{Af}$$

The formula maintains that the productivity *P* can be interpreted as the ratio between the number *h* of *hapax legomena*—words with frequency 1 in a certain text corpus—found with a certain affix and the token number *N* of the words formed with that affix in the corpus.

The concrete impact of this formula can be visualized by means of the Figure 1, in which the type/token values for three Italian suffixes forming action nouns (respectively *-mento*: *occultare* ‘to hide’ > *occultamento* ‘hiding’, *-(t)ura*: *stirare* ‘to iron’ > *stiratura* ‘ironing’, *-nza*: *credere* ‘to believe’ > *credenza* ‘belief’) and one suffix forming adverbs (*pronto* ‘ready’ > *prontamente* ‘readily’) are calculated on the basis of a large newspaper corpus containing 75 million tokens (see the details in Gaeta and Ricca 2006, 2015):

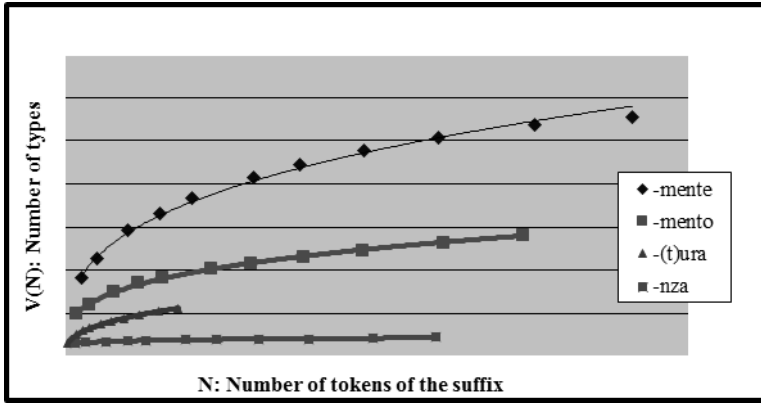


Figure 1. *La Stampa* 1996-1998:
types increasing curve as a function of N

The four curves display different slopes in correspondence of their different productivity, intended as the probability of finding a new type V after collecting N tokens of the suffix. This formula can be operationalized in at least two different ways, corresponding to different aspects of productivity. In this regard, Baayen (2009) suggests the concept of “realized productivity” ($= P(N_{\text{Max}})$) calculated at the final point of the curve, i.e. including all tokens found in the corpus with a certain affix, which refers to the level of the Lex_1 because it measures how far an affix is from saturating its domain of application. A different procedure suggested by Gaeta and Ricca (2006) and termed “expanding productivity” by Baayen (2009) ($= P(N_{V.C})$) is calculated at the same N point of the curve for all affixes, which entails that operationally one has to extract the $P(N_{V.C})$ values from different subcorpora for all affixes since the latter display different (even highly different) frequency values in the total corpus. Therefore, this procedure requires a variable-corpus approach, provided that the frequency values for the affixes remains constant throughout the corpus. The “expanding productivity” refers to the level of the Lex_2 because it compares the contribution of two affixes to the growth rate of the vocabulary in a corpus, i.e. it is a way to rank their productivities, abstracting away from the degree of frequency of the affixes in the Lex_1 . This effect is due to the fact that for each affix $P(N)$ is not constant, but is a decreasing function of N , even tending to zero when N approaches infinity. The decreasing monotonic nature of the function $P(N)$ is visualized by the curves reported in the following figure which plots the $P(N)$ values of the same four suffixes on their N values:

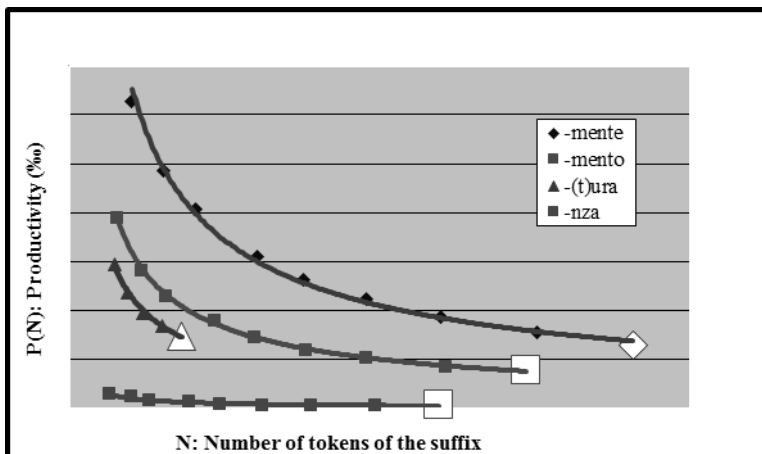


Figure 2. Productivity as a function of N

Since the “realized productivity” compares the productivity values at the end of the curves, it always implies an overestimation of the values of P for the less frequent suffixes due to the decreasing character of all $P(N)$ curves. Therefore, the “realized productivity” $P(N_{\text{Max}})$ expresses the degree of Lex_1 -entrenchment of the suffixes. On the other hand, the “expanding productivity” $P(N_{V-C})$ compares the productivity values at the same point N keeping their frequency under control and focusing on their probability of giving rise to new Lex_2 -lexemes.

That the two procedures can provide very different results is shown by cases of affixes that display very different frequency values (and therefore a different “realized productivity”) but that at the same time are expected to be similar with regard to the “expanding productivity” by virtue of a similar derivational meaning and accordingly a similar input. One good example is provided by the two Italian suffixes *-tore* and *-trice* forming respectively masculine and feminine agent nouns: *educare* ‘to educate’ > *educatore* ‘educator’/*educatrice* ‘educator:FEM’.

	$P(N_{\text{Max}}) \cdot 10^3$	$P(N_{V-C} = 23\,780) \cdot 10^3$
[V <i>-tore</i>]: $N_{\text{Max}} = 160\,142$	2.7	9.0
[V <i>-trice</i>]: $N_{\text{Max}} = 23\,780$	9.4	9.4

Table 1. Masculine and feminine agent nouns in Italian

As can be gathered by Table 1, while the “realized productivity” $P(N_{\text{Max}})$, connected with the Lex_1 , is strongly affected by the frequency values as the masculine scores three and a half time less productive than *-trice* in correspondence of their much higher token number, the values for the “expanding productivity” $P(N_{\text{V.C}})$ are substantially aligned, which mirrors quite closely their similar status at the level of the Lex_2 .

Thus, in the light of what we can gather from the phenomena of blocking and of productivity, distinguishing the Bloomfieldian Lex_1 from the Lex_2 helps us explain the substantial differences observed. In particular, far from being useless, the Lex_2 qualifies for an important theoretical construct, which plays a crucial role in the two core phenomena discussed above as well as in other cases of no less relevance such as the assumption of possible but unattested words as the input of word-formation rules. For instance, *decaffeinate* presupposes the unattested *°caffeinate* which accordingly qualifies as a pure Lex_2 -lexeme.

2.4 Lex_1 -lexemes, Lex_2 -lexemes and the place of the lexicon

Once that the assumption of the Lex_2 has been widely justified empirically, the natural question arising now concerns the relation between the Lex_1 - and the Lex_2 -lexemes. One answer which tries to avoid what has been called the rule/list fallacy (cf. Langacker 1987, 29) is provided by Goldberg’s (2006, 2013) idea of a Constructicon which we might consider to contain the Lex_1 and the Lex_2 :

Constructicon	Examples
Word	<i>Iran, another, banana</i>
Word (partially filled)	<i>pre-N, V-ing</i>
Idiom (filled)	<i>Going great guns, give the Devil his due</i>
Idiom (partially filled)	<i>Jog <someone’s> memory, <someone’s> for the asking</i>
Idiom (minimally filled): <i>The Xer the Yer</i>	<i>The more you think about it, the less you understand</i>
Ditransitive construction: Subj V Obj1 Obj2 (unfilled)	<i>He gave her a fish taco; He baked her a muffin</i>
Passive: Subj AUX VP _{PastPart} (PP _{by}) (unfilled)	<i>The armadillo was hit by a car</i>

Table 2. A fragment of the English Constructicon (cf. Goldberg 2013)

Given this far-reaching understanding of the lexicon as a Constructicon containing any sort of fully specified or almost completely unspecified pattern, we might carve out the place for morphology by exploiting the idea that the latter is basically word-based. This is in fact the position maintained by Jackendoff (2002, 125) who centers on the Constructicon consisting of our Lex₁-lexemes (“the parts of language that are listed in long-term memory”) his Tripartite Parallel Architecture which, however, does not contain any formation rule specific for morphology in neat contrast to phonology and syntax (cf. Jackendoff 2002, 2013):

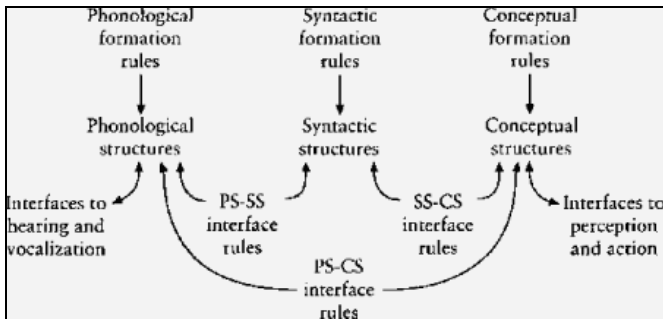


Figure 3. CAPTION NEEDED HERE

Correspondingly, Booij (2010, 11) rejects the status of a separate module for morphology insofar as the words are dealt with by word grammar which partially shares its architecture with syntax (Booij 2010, 11, emphasis added):

[t]he tripartite structure [...] of a word formation schema [...] makes clear that **morphology is not a module of grammar on a par with the phonological or the syntactic module** that deal with one aspect of linguistic structure only. Morphology is word grammar and similar to sentence grammar in its dealing with the relationships between the three kinds of information. It is only with respect to the domain of linguistic entities that morphology is different from sentence grammar since morphology has the word domain as its focus. **This architecture for morphology is the same as that for sentence grammar**, but its domain is smaller, namely that of the word.

Consistently with previous work, Booij assumes as a divide line the Lexical Integrity Principle as mentioned in section 2 above, which

basically prohibits any movement, splitting, and deletion of parts of words (Booij 2009, 97, emphasis added):

The main reason why we consider a sequence of morphemes a word is that **that sequence behaves as a cohesive unit with respect to syntactic processes**. In other words, **cohesiveness is the defining criterion for canonical wordhood**, whereas other properties such as being a listeme (a conventional expression) are clearly not to be seen as defining properties for wordhood. Hence, **if we take the notion word seriously, we might say that its defining property is cohesiveness or non-interruptability**.

This amounts to say that a sequence of morphemes qualifies as a word if it behaves as a cohesive unit with respect to syntactic processes manipulating or accessing to parts of the words. It has to be stressed that such a requirement has been often discussed in the literature in connection of what has been called here a “Lex₁-lexeme”, namely the fact of being a listeme (cf. Di Sciullo and Williams 1987) or a syntactic atom, typically resulting from a process of lexical entrenchment. These criteria for Lex₁-lexemehood are usually summarized as follows (cf. Mugdan 2015 for a recent discussion):

- i) Prosodic autonomy
- ii) Inalterability (No insertion; No extraction)
- iii) Semantic and functional autonomy (Conceptual unity)
- iv) Impenetrability (Anaphoric islandhood)
- v) Cohesion (No syntactic modification; No syntactic dislocation)

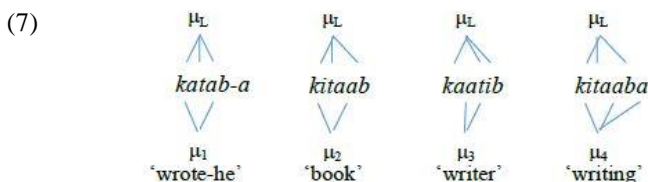
Unfortunately, no single criterion holds against empirical testing as it is not difficult to find out cases which are problematic for the definition of a lexeme. Thus, we know of cases of units that are prosodically autonomous which are not to be considered lexemes such as the German particle *ein-* found in connection with the so-called separable verbs:

- (6) a. *Ein guter Freund von mir steigt in München ein.*
‘A good friend of mine will step in in Munich’.
- b. *Ein guter Freund von mir möchte in München einsteigen.*
‘A good friend of mine wants to step in in Munich’.

Clearly, in (6a) *ein-* is prosodically autonomous insofar as it occurs quite far from the verb and it bears the main sentence stress. On the other hand, it cannot be analyzed as a lexeme as it does not occur outside of the verb *einsteigen* (and it also needs to be combined with it when this appears in a

non-finite form (6b)), although it is etymologically and semantically connected with the preposition *in*¹. Even worse, there is already in German a lexeme *ein*, namely the indefinite article.

As for inalterability, this is normally discussed in connection with the prohibition of insertions and extractions. Accordingly, one should not find any case of apophony or non-concatenative morphology like those massively found in Arabic:



As is well known, the word results from the intertwining of the consonantal backbone providing the basic lexical content (in our case: the tri-letteral root *k-t-b* means [WRITE]) and of the vocalic flesh which specifies the morphological meanings: μ_1 = [VERB PAST], μ_2 = [RESULT NOUN], μ_3 = [AGENT NOUN], μ_4 = [ABSTRACT NOUN]. In such comb-shaped morphology, in which the single consonants and vowels are intertwined like the teeth of two combs, the identity of the lexemes results from abstract patterns where the idea of an insertion is simply inadequate to capture their essence. On the other hand, the criterion of extraction is falsified by the property of separability illustrated by the German particle *ein-* in (6a) above.

The criterion of semantic and functional autonomy, which is also referred to as “conceptual unity”, pertains to the alleged unity of the meaning of a lexeme in contrast with the compositional nature of the meaning of a syntactic unit. Since also the meaning of complex lexemes can be compositional, I rather prefer to speak of onomasiological stability

¹ Theoretically, one might also claim that *ein-* is the form assumed by the preposition *in* when it forms a compound with a verb, but this basically amounts to saying that *ein-* is a prefix of the verb, because it is a bound morpheme. Moreover, the form *in* is normally found in verbal compounds resulting from univerbation: *ineinanderfließen* ‘to merge into one another’, *instandbesetzen* ‘to squat and refurbish’. Finally, in many cases the (basically spatial) prepositional meaning is not easy to reconstruct in the verb (e.g. *einhalten* ‘to hold, observe the rules’, *einkaufen* ‘to do the shopping’, *einklagen* ‘to sue sb. for sth.’), which neatly contrasts with the widespread compositionality observed in true compounds while the meaning of the affixes is usually much more difficult to identify precisely.

of a lexical unit with respect to the allegedly unstable meanings of the syntactic units (cf. Gaeta 2015b). That such alleged onomasiological stability cannot be of any help to distinguish between lexical and syntactic units is shown by the following examples:

- (8) a. *Rotwein* vs. *Rotes Kreuz*
‘red wine’ ‘Red Cross’
b. *golletto salva-Roma,* *voto salva-Berlusconi*
‘FC-Rome-saving little goal’ ‘Berlusconi-saving vote’
c. *La mia lezionecina è durata ben due ore.*
‘My lesson:DIM lasted two good hours’.

In (8a), *Rotwein* and *Rotes Kreuz* are not essentially different with regard to their semantic autonomy, although the former is clearly a compound and the latter a syntactic unit of German. On the other hand, the Italian VN-pattern for compounds gives normally rise to examples like those in (8b) which cannot be really considered to display any conceptual unity, and rather look like shortened relative sentences. Finally, the morphopragmatic value of the Italian diminutive in (8c) does not refer to any shortened period of time but is rather used by the speaker to hide herself behind a certain degree of real or fictive modesty (the so-called *diminutivum modestum*, cf. Gaeta 2015c). It goes without saying that such a morphopragmatic value is rather difficult to be lexicalized.

Impenetrability usually refers to the property of being an anaphoric island typical of lexemes in contrast with syntactic units. However, it is not difficult to find examples showing that Italian VN-compounds can be penetrated by anaphors, while syntactic units stabilized in the lexicon (our Lex₁-lexemes) are completely impenetrable:

- (6) a. *collanina* “*trasgressiva*” [*acchiappa*_i[*talenti*]_i] *tra i quali*_i *ha*
figurato anche il primo Ammanniti
 ‘*talent_i-catching raffish little necklace, among whom_i there was*
also the young Ammanniti’
 b. **Ho una camicia blu di Prussia*_i, *ma non ci*_i *vado mai*.
 ‘I have a Prussia_in blue shirt, but I never go there_i’.

The same is true for Booij's crucial criterion distinguishing morphology and syntax, namely cohesion. On the one hand, both Lex₁-lexemes and compounds do not allow any insertion of modifiers with respect to syntactic units (10a)-(10c) or any extraction and anaphoric reference in contrast with typical left dislocations (10d)-(10e):

- (10) a. *[[tappetino bianco] da bagno]/[[tappetino [da bagno]] bianco]*
 ‘white bath mat’
 b. **[ferro [giallo] da stiro]/[[ferro da stiro] giallo]*
 ‘yellow flatiron’
 c. **[capo [siciliano] mafia]/[[capo mafia] siciliano]*
 ‘Sicilian mafia boss’
 d. *di tappetini_i ne_i ho acquistati due _i da bagno.*
 ‘I bought two bath mats’.
 e. **di ferri_i ne_i ho comprati due _i da stiro.*
 ‘I bought two flatirons’.
 f. **di capi_i ne_i ho arrestati tre _i mafia.*
 ‘I arrested three mafia bosses’.

In neat contrast with what should be expected on the basis of the Lexical Integrity Principle, compounds and Lex₁-lexemes behave in the same way as for the possibility of being manipulated by syntactic rules. In other words, all the criteria listed as i)-v) above, including cohesion, are only able to identify Lex₁-lexemes with respect to syntactic units, i.e. the assess the Lex₁-status of any unit independently of its morphological or syntactic nature. Since it turns out that those criteria are only able to identify Lex₁-lexemes, it is not surprising to observe that in some linguistic tradition compounds are not distinguished in a principled way from frozen phrases (Apothéloz 2002, 18):

La composition est un mode de formation incluant un éventail relativement large de phénomènes, entre lesquels les linguistes ont souvent cherché à établir des distinctions. Le point commun de ces différentes formations est sans aucun doute le figement. Généralement, on entend par “composition” la construction d’une unité lexicale complexe au moyen d’un morphème grammatical non-affixal et d’un morphème lexical (*sans-abri*, *arrière-boutique*), ou d’au moins deux morphèmes lexicaux libres ou liés, pouvant donc eux-mêmes servir de base à une dérivation (*chou-fleur*, *grand-père*, *lave-vaisselle*, *compte-gouttes*, *bibliophile*, *ludothèque*), les morphèmes libres pouvant être accompagnés d’un ou plusieurs morphèmes non lexicaux (*arc-en-ciel*, *eau-de-vie*)².

² “Compounding is a way of formation including a relatively large number of phenomena, among which the linguists have often tried to establish some distinctions. The common point of these different formations is undoubtedly frozenness. Generally, one understands under compounding the construction of a complex lexical unit by means of a non-affixal grammatical morpheme and of a lexical morpheme (*sans-abri* ‘homeless, lit. without-shelter’, *arrière-boutique* ‘backshop’), or by means of two free or bound lexical morphemes which can also

In this way, the distinction between the Lex₁ and the Lex₂ is completely blurred insofar as there does not seem to exist any principle independent of Lex₁ which can vouch for the existence of Lex₂. This is so because no apparent distinction can be made between constructions within the Constructicon. If we only rely on the criteria used for defining canonical wordhood in Booij's sense, then we are only able to identify partially or entirely fixed units, i.e. Lex₁-lexemes, including multi-word expressions: [*pre-N*], [*The X-er the Y-er*], etc. Since such Lex₁-units encompass several different objects, simply speaking of the word level is insufficient to keep Lex₂-lexemes (which are expected to result from morphology proper) from Lex₁-lexemes. Unless we can mention independent principles able to identify a consistent lexical morphology explicitly referring to the Lex₂. In the next section, I will suggest how to deal with this question keeping the Lex₁ and the Lex₂ apart.

3. Morphology between the Lex₁ and the Lex₂

Given the basic unreliability of the definition of the canonical wordhood to identify the restricted domain of morphology with regard to syntax, one might be tempted to reject any separate status and rather treat morphology and syntax “[...] as different scales of phrasal syntax with different behavior, much as different scales of phonology such as phonological words and intonational phrases have somewhat different principles” (Jackendoff 2002, 129). As observed in section 2 above, this is the line adopted by Distributed Morphology, in which the concept of lexeme does not play any valuable role. In this section, I will try to explore the opposite view, namely the idea of a morphology as a separate module with respect to syntax insofar as it can be explicitly shown to display independent organizational principles. The concept of module should be interpreted in what follows in rather broad terms as a component displaying distinct and autonomous properties which on the one hand interact with other modules while on the other they are not strictly separate and rather form a continuum mirroring the complexity of the phenomena. The reference to the continuum is not meant to introduce in the theory an escape strategy

serve as base for a derivation (*chou-fleur* ‘cauliflower’, *grand-père* ‘grandfather’, *lave-vaisselle* ‘dishwasher, lit. wash-crockery’, *compte-gouttes* ‘dropper, lit. count-drops’, *bibliophile* ‘bibliophile’, *ludothèque* ‘games library’), while the free morphemes can be accompanied by one or more non-lexical morphemes (*arc-en-ciel* ‘rainbow, lit. arc-in-sky’, *eau-de-vie* ‘brandy, water-of-life’)” (my translation).

able to accommodate possible exceptions with the effect of dissolving the peculiar character of the component in an indistinct series of phenomena where a divide line is concretely impossible to draw. Notice that this is the solution adopted by Apothéloz for French compounding, where it is allegedly impossible to distinguish between what comes out from morphology proper and what results from the freezing of syntactic units. Instead, I will adopt the idea, quite common in Prototype Theory (cf. Geeraerts 2010, that a category, in our case morphology, can be properly defined by means of explicit properties profiling the so-called degree of membership, while the prototypical effects arise with regard to the “degree of representativity” (cf. Geeraerts 2010, 191). In this light, elaborating on Corbin (1997), I will adopt as a defining criterion for identifying morphological operations the M(orphological)-Principle (cf. Gaeta 2015b):

M-Principle: Lexeme-formation operates at the level of X_0 and cannot be arguably reduced to syntax.

This criterion contains a positive side, namely the reference to the level of X_0 as basic operative domain, and negative side which needs to be spelt out in three different corollaries:

- i) **C-1:** The sequence cannot be straightforwardly generated by syntax.
- ii) **C-2:** The phonology of the sequence is different from that of a syntactic unit.
- iii) **C-3:** The phonology of the sequence is different from that of a syntactic unit.

The level of X_0 basically refers to Aronoff's (1994, 11) definition of the lexeme intended as “[...] a (potential or actual) member of a major lexical category, having both form and meaning but being neither, *and existing outside of any particular syntactic context*” (my emphasis). Morphology, and for our purposes word-formation, prototypically consists in giving rise to new lexemes, while inflectional morphology prototypically provides single forms of a lexeme to be used in a given syntactic environment.

The three corollaries focus on concrete properties of the expressions which help us distinguish morphological from syntactic units. Here, I will only discuss the first corollary while the others are treated in details in Gaeta (2015b). Accordingly, C-1 allows us to structurally distinguish a morphological unit as non-generatable by syntax. Such concrete properties generally have language-specific character and cannot be therefore

extended a priori cross-linguistically. To make one concrete example, syntax cannot generate NN sequences presenting an internal relationship of a subordinative type in French, while this is the case in German where the subordinate noun is overtly marked:

- (71) a. *Le chapeau *(de) Pierre* / *Der Hut Peters*
 ‘The hat of Peter’
 b. *Une petite gorgée *(de) vin* / *Ein Schlug Wein*
 ‘A slug of wine’
 c. *Le timbre (*de) poste* / *Die Briefmarke*
 ‘The postage stamp’
 d. *Les lunettes (de) soleil* / *Die Sonnenbrille*
 ‘The sunglasses’

However, the picture in German is much more complex than the rough picture given here as, for instance, the subordinative relation is only possible with proper nouns, which are overtly marked by the gender-independent suffix *-s* (cf. *Der Hut Annas* ‘the hat of Anna’), while common nouns normally require a more complex structure usually opened by a prepositional phrase or a full NP including an overt determiner: *der Hut (von) der Wittwe* ‘the hat of the widow’. One particular case is given by (11b) in which the NN sequence *Schlug Wein* occurs, which might be potentially interpreted as non-syntactic, and therefore morphological. That this cannot be the case is shown by the comparison with typical NN-compounds which display modifier-head structure and give rise to a clear contrast with the syntactic structure of (11b) in which the head lies at the left side as shown by (11c) and also by pairs like *ein Glas Wasser* ‘a glass of water’/*ein Wasserglas* ‘a glass for water’. Such a syntactic structure can only be headed by nouns referring to containers or more in general quantifying expressions while the modifier typically is a liquid or a quantifiable substance. As shown by the contrast between French and German, morphological (left- vs. right-headedness) and syntactic (occurrence of prepositions) structure helps us distinguish between what is morphologically and what is syntactically constructed. Notice that in (7d) the optionality of the preposition in French configures a more complex situation which has been traditionally answered by making reference to the syntactic origin of the unit. For instance, Fradin (2009, 433) observes for French that “[m]any such expressions generally have both structures, the one with the preposition being the oldest one”. While this might be true, it does not solve our problem of where to put these units, whether in the syntax or in the morphology. In fact, two interpretations are possible:

either an entirely syntactic one as currently done in Distributed Morphology (cf. section 2 above), or a diachronic one. The latter interpretation follows from Apothéloz' approach mentioned above, which considers compounds as basically resulting from a process of freezing. Thus, the alternation between the type with or without preposition is basically to be seen in purely diachronic terms. This implies that a unit like *timbre-poste* which cannot be synchronically derived from a syntactic unit has to be seen as Lex₁-lexeme on a par with *arc-en-ciel*, *eau-de-vie*, etc. That this cannot be true is shown by the synchronic productivity of this pattern as has been investigated carefully by Arnaud (2003), who has reconstructed the development of the synchronic pattern starting from its diachronic origin in frozen NPs until its subsequent generalization in purely morphological terms in the last century.

A similar and in a way subtler case is given by the following Italian examples, to a large extent paralleled by French examples on their right, and in which respectively a non-argumental and an argumental head occurs:

- (12) a. [N_i N_j]_i *centro congressi* / *timbre-poste*
 'congress center' 'postage stamp'
 b. [N_i N_j]_i *responsabile donne* / *responsable femmes*
 'consultant on women's issues'

Similar to the French case, the label "reduced phrases" has been suggested also for these Italian examples (cf. Scalise and Bisetto 2008, 138). Again, while this label may capture the diachronic origin of these constructions, it is not clear what they mean from a synchronic point of view. If we stay in the Lexicalist field, as Scalise and Bisetto actually do, the process of reduction can only mean deletion, which, if not properly constrained, paves the way for the unrestricted effects of the Transformationalist view criticized in the early 1970s. One way of constraining these unwelcome effects comes from a suggestion of Baroni, Guevara and Zamparelli (2009) who account for these reduction processes in terms of the generalization of an operation typically characterizing the headlines style in which the syntactic expressions are reduced for brevity basically via the suppression of grammatical morphemes. While this might surely contribute to the diffusion of these patterns, especially with regard to their brevity, two facts speak against this view (cf. Gaeta 2015b). First, there is no strict correspondence between the type of reduced phrases resulting from headlines-like reduction and compounds which are structurally possible in Italian:

- (8) a. *Caos M5S dopo ira Grillo*
 Chaos M5S after anger Grillo
 ‘Chaos in M5S after Grillo’s anger’.
- b. *Niente attacco se consegna armi gas*
 Nothing attack if delivery weapons gas
 ‘No attack if the gas weapons are delivered’.
- c. *La gestione del [problema armi chimiche] in Siria, ricorda molto quello che accadde nel 2003 in Iraq.*
 ‘The management of the question of the chemical arms in Syria reminds closely what happened in 2003 in Iraq’.

While the reduced phrases *ira Grillo* and *armi gas* are found in typical newspapers’ headlines like (13a)-(13b), they are not possible Italian compounds as their derivational meanings, respectively referring to the material (*??armi gas*) and to the experiencer (*??ira Grillo*), are usually unproductive: cf. *??tavolo legno* ‘wooden table’ and *??rabbia gente* ‘wrath of the people’, in contrast for instance with their German correspondences *Holztisch* and *Volkszorn*. Second, instead of the headline expression *armi gas*, another expression is found in compounds which consists of the stabilized unit, i.e. the Lex₁-lexeme, *armi chimiche* ‘chemical arms’ (13c). This draws us back to the examples in (127), as they normally show a large typology of cases in which either the head or the modifier consists of Lex₁-lexemes:

- (9) [N_i [N_j]_i *centro congressi internazionali*
 ‘international congress center’
 [[N_i N_j]_i *responsabile nazionale donne*
 ‘national consultant on women’s issues’

Notice that Lex₁-lexeme occurring in these compounds does not consist of a full NP but of a [N_i deprived of a determiner, and therefore not immediately licensed in syntax: **centro i congressi internazionali*. In this way, these structures comply with the M-Principle and the C-1, being not generatable by syntax. Such a conclusion allows us to formulate the Lexicality of the Input Principle (= LIP, cf. Gaeta 2015b):

LIP: Lexeme formation is based both on Lex₁- and Lex₂-lexemes.

The LIP takes into account the different nature of the Lex₁ and of the Lex₂, as on the one hand it accounts for complex words formed on the basis of possible but unattested words (i.e. pure Lex₂-lexemes): *to decaffeinate* <

°to *caffeinare*. On the other, it is not biased against complex units being stabilized in the Lex₁ which, according to Booij's strict Lexicalist view, should not mix with lexeme formation proper. In fact, the occurrence of []_N, i.e. syntactic units without specifier, in any sort of lexeme formation pattern is a pervasive phenomenon in Italian, and in other Romance languages as well (cf. Gaeta 2015b), provided that the unit []_N is a Lex₁-lexeme:

(10) a. AdjN Compounds

[Adj _i N _j] _i	<i>grigio perla</i> 'pearl-grey'
[Adj _i [] _{Nj}] _i	<i>grigio fumo di Londra</i> 'London smoke grey'

b. VN Compounds

[V N _i] _j	<i>portacenere</i> 'ashtray'
[V [] _{Ni}] _{Nj}	<i>porta carta igienica</i> 'toilet-paper holder'

c. Prefixation (included in a compound)

[N _i [<i>anti</i> - N _j]] _i	<i>allarme antifurto</i> 'antitheft alarm'
[N _i [<i>anti</i> - [] _N]] _i	<i>marcia antimoschea di Lodi</i> 'anti-mosque of Lodi march'

d. Suffixation

[N _i -ista _{Nj}] _j	<i>macchinista</i> 'engine driver'
[[] _{Ni} -ista _{Nj}] _j	<i>larghintesista</i> 'supporter of a coalition government, lit. broad-agreements-ist'

e. Prefixation and Suffixation

[<i>anti</i> - [[] _{Ni} -ista _{Nj}]] _j	<i>antilarghintesista</i> 'opponent of a coalition government'
--	---

Although examples like these can be easily multiplied, it should not be forgotten that the big bunch of lexeme formation is based on words rather than syntactic units, which is the reason why one normally assumes that words are the prototypical input of lexeme formation patterns (cf. Dressler 1988). In this connection, it is important to mention Corbin's (1997, 59) observation, who also opposes the concept of lexicalization to the process of lexeme formation:

la morphologie a davantage vocation à construire des unités lexicales que les autres composants de la grammaire. Mais d'une part, [...] ses produits ne sont pas automatiquement lexicalisés, d'autre part elle n'est pas la source exclusive de production des unités lexicalisées³.

On this basis, we can define lexicalization or lexical entrenchment as the process of entering into the Lex₁, i.e. becoming a Lex₁-lexeme. At the same time, this highlights a general tendency for Lex₂-lexemes to become Lex₁-lexemes. In this regard, examples are easy to find. The so-called nonce-formations typically are Lex₂-lexemes which are not (yet) established as Lex₁-lexemes. On the other hand, there are Lex₂-lexemes which are rather unlikely to become Lex₁-lexemes. This is typically true of Lex₂-lexemes whose derivational meaning is rather abstract in the sense of scarcely referential. The usage of diminutive suffixes with morphopragmatic value typically gives rise to nonce-formations, often implying the extension of lexeme-formation pattern beyond its normal domain. For instance, in this Spanish example the diminutive suffix appears on the gerund *calland-ito* providing the assertion with a non-serious character (cf. Gaeta 2015c):

- (16) *Debíamos de acercarnos callandito–sugería Fernando.*
 ‘We have to get close keeping quiet–Fernando suggested’.

Predictably, it is highly improbable for *callandito*, as well as for other cases of diminutive suffixes with a morphopragmatic value as that in (9c) above, to be entrenched as a Lex₁-lexeme. At any rate, while the tendency Lex₂-lexeme > Lex₁-lexeme is fairly common, the opposite phenomenon is marginal, namely the passage Lex₁-lexeme > Lex₂-lexeme: under this type, we can understand the process of folk etymology whereby speakers remotivate a Lex₁-lexeme by attributing a new morphological structure which is not justified etymologically. One such case is provided by the German verb *hantieren* ‘to handle, manipulate’ which goes back to a Middle French verb *hânter* meaning ‘to haunt’. Landed in Middle German in the meaning ‘to have a deal’, the verb was later reanalyzed as derived from *Hand* ‘hand’ in combination with the loan suffix *-ieren*, also found

³ “Morphology has more vocation to build lexical units than the other components of grammar. But on the one hand, [...] its products are not automatically lexicalized, and on the other it is not the exclusive source of production of lexicalized units” (my translation).

with native bases as in *gastieren* ‘to guest’, *hausieren* ‘to hawk’, etc. Correspondingly, a new meaning was attributed to the verb including that of the lexical base.

4. The explanatory force of the LIP

Even if lexicalization might at first sight appear as an inertial force, and has been indeed approached in this way for decades⁴, we will see in this section that its interaction with the Lex₂ provides us the key to understanding a number of phenomena which have been traditionally considered marginal or even irrelevant for the issue of the interaction of morphology and syntax.

Traditionally, the process of lexicalization, or in our terms of becoming a Lex₁-lexeme, has been connected with three phenomena, all originating from the epiphenomenal frequency increase: this is accompanied by: a) a meaning extension leading to enlarged contexts of use, which favors the rise of b) meaning idiomaticity, and of c) morphophonological opacity. While it is not obvious to assess whether it is the frequency increase that leads to enlarged contexts of use and therefore to meaning extension or vice-versa, namely whether it is meaning extension that forces an enlarged distribution and therefore a frequency increase, idiomaticity and opacity can be safely taken as reliable indicators to signalize the landing in the Lex₁. This is probably the reason why, in the absence of other empirical sources, they have been traditionally focused on in order to assess the Lex₁-status of an expression, which also explains the focus on lexicalization as an inertial force typical of the traditional approach.

However, since text corpora have become largely available, the picture has dramatically changed fostering the development of usage-based approaches (cf. Barlow and Kemmer 2000) which entirely rely on “behavioral” factors such as frequency and degree of frozenness or collocational strength (cf. Gries 2013) in order to figure out in dynamic terms an “Emergent Grammar” of a language (cf. Hopper 1987). In this connection, it should not be forgotten that these ideas lay behind the

⁴ On this subject see for instance Bauer (1983, 50): “[l]exicalization, as it has been described here, is essentially a diachronic process, but the traces it leaves in the form of lexicalized lexemes have to be dealt with in a synchronic grammar” and twenty years later Plag (2003, 91, original emphasis): “[a]part from the compositional meaning just described, many *-ity* derivatives are **lexicalized**, i.e. they have become permanently incorporated into the mental lexicons of speakers, thereby often adopting idiosyncratic meanings”.

concept of the Constructicon as it has been sketched in section 2.4 above. In this section, I will illustrate the role of the Lex₁ in its interaction with the Lex₂ in order to account for two phenomena which have been either considered marginal or discarded as irrelevant.

The first case is given by phrasal compounds which, in spite of their severe violation of the Lexical Integrity Principle, are traditionally left apart in the theoretical discussion. For instance, according to Scalise (1994, 142), Italian phrasal compounds like *ragazze casa e chiesa* ‘home-and-church-girls’ are rather marginal and normally arise in jocular situations basically by means of shortening: *ragazze (che stanno sempre in) casa e chiesa* ‘girls who always are either at home or in the church’ (cf. Gaeta 2003, 2006, 2015b for a different view). Clearly, in a language like German in which compounding plays such a big role, phrasal compounds display a much wider typology (cf. Lawrenz 1996, Meibauer 2003, 2007):

- (1711) a. Idioms: *eine/die Kopf-durch-die-Wand-Strategie*
 ‘a/the head-through-the-wall strategy’
 b. Clichés: *eine/die Ich-traue-mich-nicht-Hommage*
 ‘a/the I-don’t-dare homage’
 c. Titles: *ein/das Romeo-und-Julia-Gefühl*
 ‘a/the Romeo-and-Juliet feeling’
 d. Quotes: *ein/der Ich-bin-ein-Berliner-Auftritt*
 ‘a/the *Ich-bin-ein-Berliner* performance’
 e. Binomials: *die Pfeffer-und-Salz-Haare*
 ‘the pepper-and-salt hair’
 f. Loan phrases: *ein/das Fast-Food-Kino*
 ‘a/the fast-food cinema’
 g. Quantity expressions: *eine/die 180-Grad-Wende*
 ‘a/the 180-degrees turn’

Three important properties have been specified. As in normal compounding, they are right-headed and stressed on the modifiers, which retain their phrase accent: *Kopf-durch-die-Wand-Strategie*, *Ich-traue-mich-nicht-Hommage*, etc. Notice in this regard pairs like *Vórort-Tarif* ‘suburban tariff’ vs. *Vor-Órt-Tarif* ‘local, lit. before-place, tariff’. Moreover, they are mostly headed by nouns, although (derived) adjectival or participial heads are not excluded: *Blut-und-Boden-mystisch* ‘blood-and-soil mystical’, *Fünf-Stufen-intergriertes Filtersystem* ‘five-steps integrated filtering system’. Finally, modifiers cannot be full NP including a determiner, but only a []_N, similarly to the Italian compounds in (9):

- (128) a. *ein/der (*die-)graue-Schläfe-Effekt*
 ‘a/the (*the-)grey-temple effect’
 b. *eine/die der-schöne-Rheingau-Laberei*
 ‘a/the the-beautiful-Rheingau talk’

A partial exception is constituted by those cases in which the determiner is already part of the frozen unit, as in (128b). This brings us to the main point which underlies the occurrence of phrasal compounds: their modifiers generally consist of Lex₁-lexemes. This does not mean that they must be treated as “quotations” which are arbitrarily inserted into the morphological structure as suggested for instance by Wiese (1996), because the quotative value accounts for only one type, while the typology is much wider and comprises in fact any sort of stabilized expression (from typical idioms to quantized expressions) which can enter the Lex₁, as illustrated in (17) above. While these Lex₁-lexemes constitute the absolute majority of modifiers of phrasal compounds (cf. Lawrenz 1996), deviations from this general tendency are possible, insofar as expressions made out of the blue which are clearly not established as such can enter phrasal compounds:

- (19) a. *Teenager-finden-sich-und-ihre-Liebe-Prinzip*
 ‘Teenagers-find-themselves-and-their love principle’
 b. *Zap-und-weg-Fernsehzeiten*
 ‘Zap-and-away TV times’

However, they must be interpretable as established, at least for the purposes of the speech situation, with the help of pragmatic principles. In particular, Meibauer (2007) draws attention to the fact that the particular expressive effect of phrasal compounds results from a conflict between two principles regulating the speakers’ interaction during a concrete speech situation: the Principle of Informativeness (= PI) favoring minimal informativity: “Say as little as necessary” and the Principle of Quantity (= PQ) enhancing maximal informativity: “Do not provide a statement that is informationally weaker than your knowledge of the world allows”. While normal compounds usually observe the PI because they require the enrichment of a minimal and underdetermined structure from the listener’s side, phrasal compounds follow the PQ (Meibauer 2007, 248):

[...] if a speaker knows that there is a lexicalized construction, e.g. a title or a cliché, or a quotation that enhances informativity, then he should use it. This exactly corresponds to the requirement “select the informationally strongest paradigmatic alternate that is consistent with the facts”.

This gives often rise to an ironic effect because compounds are not normally expected to contain such long chunks of information in neat violation of the PI⁵. The ironic effect can be restrained by the Lex₁-status of the phrasal expression which provides the phrasal compounds with prestige as a consequence of its widespread diffusion. On the other hand, if the speaker intentionally makes use of a novel phrasal compound favoring the PQ, the focus will be on the violation of the PI clearly enhancing the ironic effect since the phrasal expression is only contextually (i.e. by the speaker himself) provided with prestige, and this “[...] leads to the integration of a phrase into word structure” (Meibauer 2007, 248).

The second case showing the relevance of the LIP is even more intricate because it shows the effect of a typical syntactic phenomenon, i.e. coordination reduction, across compounds and phrases. Notice that in German the reduction of coordinated complex words is normally observed between compounds (*Herbst- und Frühlingsblumen* ‘autumn and spring flowers’) as well as between certain types of suffixations (*Gewerk- und Genossenschaften* ‘unions and cooperatives’) which behave like compounds in several respects, as for instance their lexical base appears in the typical compound form found for instance in *Genossengruppe* ‘group of comrades’, and in general with prefixation (*Ver- und Entsorgung* ‘provision and removal’, *Auf- und Abstieg* ‘rise and descent’, etc.). In a very detailed investigation, Askedal (2005) compares this phenomenon in three Germanic languages: German, English and Norwegian. With the focus on German, in which Askedal is able to identify about four hundred examples, several types of increasing complexity occur:

- (20) 2 members: *kontextuelle und Transformationsregeln*
‘contextual and transformation rules’
- 3 members: *personale, situative and Genrestile*
‘personal, situational and genre styles’
- 4 members: *lokale-, temporale, Modal- und Satzadverbiale*
‘local, temporal, modal and sentence adverbials’
- 5 members: *in finaler, koinzidenter, Subjekts-, Objekts- sowie sowie der rein semantischen Funktion des “Präsentats”*
‘in final, coincident, subject, object as well as in the purely semantic function of the presented’

⁵ Clearly, phrasal compounds are not jokes, but similarly to the jokes their ironic effect seems to do with incongruity, a fundamental notion that is for a general theory of humor: “Incongruity on the word level means that it is unusual to combine a phrasal meaning with a word meaning” (Meibauer 2007, 249).

6 members: *organisatorische, geographische, energiewirtschaftliche, planerische* sowie *Problemlösungs- und Berechnungsschema*
 ‘organizational, geographic, energy-efficient, planning-related as well as problem-solving and calculation formula’

Clearly, most types contain two or at least three members, while the others are much rarer. Nevertheless, they offer a wide typology in which compounds and phrases are intertwined insofar as either the phrasal head: *kontextuelle* und *Transformations[regeln]*; or the compound head is gapped: *in Infinitiv- und übergeordneten [Sätzen]*, although the former case is by far the most frequent. Moreover, the phenomenon involves an adjective coordinated with a compound displaying any sort of modifier:

- (21) Nouns: *kontextuelle und Transformationsregeln*
 ‘contextual and transformation rules’
- Adjectives: *öffentliche und Privatmittel*
 ‘public and private means’
- Verbs: *ein Lehr- oder sonst ein geistliches Amt*
 ‘a teaching or rather a spiritual office’
- Prepositions: *bei den abhängigen oder Neben-Sätzen*
 ‘in the dependent or subordinate sentences’
- Particles: *des lauten oder Vorlesens*
 ‘of the loud or reading out’
- Prefixes: *eine erste oder Ursprache*
 ‘a first or protolanguage’
Un- bzw. weniger Markiertheit
 ‘of un- or less markedness’
- Clippings: *in der heimischen ebenso wie der DaF-Grammatik*
 ‘in the native as well as in the foreign language grammar’
- Metalinguistic expressions: *in reinen und als-Appositionen*
 ‘in pure and als-appositions’
- Coordinated phrases: *das subalterne oder Teil-Ganzes-Verhältnis*
 ‘the subordinate or part-whole-relation’
- Loan phrases: *alle nationalen und Federal-Reserve-Banken*
 ‘all national and Federal-Reserve banks’

What all the phrasal units involved in such coordination reductions have in common is the same property: they are part of the Lex₁, i.e. they are Lex₁-

lexemes. This is actually the conclusion reached by Askedal (2005, 20, emphasis added):

Das Zustandekommen intrakategorieller Homogenität in koordinativen Verknüpfungen ist offensichtlich **von lexikalischen Bedingungen** abhängig. Dabei handelt es sich um Lexikalisierung auf unterschiedlichen Ebenen der Grammatik. Von Bedeutung ist zum einen die Verfügbarkeit von Adjektivlexemen, zum anderen aber auch **Usualisierung und Terminologisierung sowohl von Komposita wie auch von NPs** aus attributivem Adjektiv und Substantiv. Semantisch einschlägige Lexikalisierung ist nicht nur eine Eigenschaft von Einzelwort-Lexemen, sondern auch von mehrwortigen Phrasen bzw. Syntagmen.”⁶

This means that the intertwining of phrasal expressions and compounds, which constitutes a systematic violation of the Lexical Integrity Principle, seems to be made possible by their common Lex₁-status. Far from being surprising, in our view this conclusion is expected on the basis of the LIP, as the lexicality of both phrasal units and compounds makes it possible to align them both as Lex₁-lexemes on the same syntactic string, which is subsequently reduced if they share the head noun N_i:

$$[[[X' \text{ } \cancel{N_i}]_{N_i}^0 / [\text{Adj } \cancel{N_i}]_{N_i} \text{ PART}_{\text{Coord}} [[X' \text{ } N_i]_{N_i}^0 / [\text{Adj } N_i^0]_{N_i}]_{NP}$$

In an attempt to rescue the Lexical Integrity Principle, Booij (1985) provides an explanation of this phenomenon in purely prosodic terms which is based on a neat distinction between the morphosyntactic level and the purely prosodic one. This is justified by the allegedly primary role played by the purely prosodic constituency of the deleted constituents, which is taken to support Selkirk's (1982, 70) “Autonomy of Word Structure Condition”, which states that “[n]o deletion or movement transformation may involve categories of both W[ord]-structure and S[yntactic]-structure”. This analysis is based on the idea that gapping is a prosodic phenomenon which is sensible only to “[...] an independent prosodic structure which is not necessarily isomorphic to morphological/syntactic structure” (Booij 1985, 156). While the prosodic

⁶ “The occurrence of intracategorical homogeneity in coordinative conjunctions clearly depends on lexical conditions. In fact, it is a question of lexicalization at different grammar levels. Of significance is the availability of adjectival lexemes on the one hand and on the other also the usualization and terminologization of compounds as well as of NPs formed by adjectives and substantives. Semantically-conditioned lexicalization is not only a property of single-word lexemes but also of multiword expressions and phrases” (my translation).

relevance of gapping is undisputable, it seems to me that restricting its relevance to prosodic constituents, independently of their morphological or syntactic status, simply bypasses the question without providing an answer to the fact that morphological and syntactic structures are at stake here.

5. Conclusion

It is now time to sum up the main findings of this paper. First, it has been laid down that the Constructicon allows us to overcome the rule-list fallacy encapsulated in the traditional approaches to the lexicon as a store. The Constructicon has to be viewed as emerging from the Lex₁ (in the very specific sense of Hopper's (1987) Emergent Grammar) and represents the interface between at least two independent modules or tiers of the grammar, respectively related to the syntactic constructions and to the morphological constructions, the Lex₂. They are independent insofar as the latter cannot be arguably reduced to the former. This is supplied by the M-Principle in conjunction with its corollaries. Along with other principles which could not be discussed here such as the Filter-Principle (cf. Gaeta 2015b), one important principle which has a strong impact on the concrete instantiation of the Lex₂ is the LIP. This again emphasizes the crucial role of the Lex₁ which continuously feeds the Lex₂ on the one side and the syntactic constructions such as the coordination (reduction) on the other, giving rise to an apparently chaotic tangle as shown by Italian compounds, German phrasal compounds and German coordination reduction. This apparent chaos might at first sight speak in favor of a reductionist view such as that espoused by Distributed Morphology. The latter is however unable to account for the crucial role played by the Lex₁-status of the units involved into these phenomena. On the other hand, the approach invoked by Booij which only relies on the Lex₁-properties of the word as they are declined by the Lexical Integrity Principle turns also out to be reductionist insofar as it is unable to capture the parallel behavior played by the different sorts of the Lex₁-expressions within the Lex₂-patterns. In this light, it is not surprising that Booij sees no reason to assume a separate morphological module as the only apparent difference is the allegedly word-based nature of morphology with regard to syntax. Notice that the tripartite view illustrated in Figure 3 above in which morphology is not separate from syntax is adopted by Jackendoff (2002, 129) only for convenience while its superiority over the alternative

hypothesis assuming the independence of morphology and syntax remains to be empirically tested:

[t]hese differences suggest that phrasal syntax and morphosyntax might be regarded as semi-autonomous tiers with related but not identical organizing principles. Alternatively, they might be treated as different scales of phrasal syntax with different behavior, much as different scales of phonology such as phonological words and intonational phrases have somewhat different principles. Working out even a sketch of these alternatives is, however, beyond the scope of the present work.

If, however, one firmly believes that there is ground to keep morphological constructions distinct from the syntactic ones and that both modules or tiers share the same Constructicon, then we might easily imagine a quadripartite view in which the Constructicon appears as the usage-based interface of the (semi-)autonomous modules of the grammar:

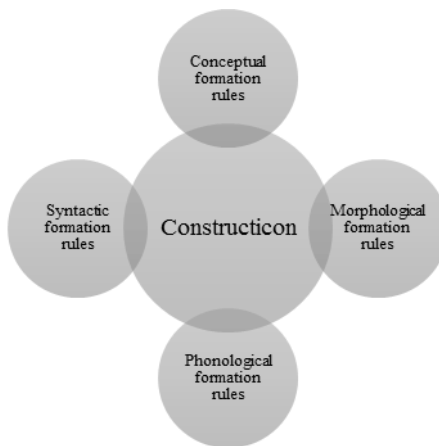


Figure 4. CAPTION NEEDED HERE

As direct interface of the autonomous modules, the Constructicon makes it superfluous to have direct connections among the single modules, which nicely corresponds to the idea that it is within the constructions, as primary objects of linguistic analysis, that the single modules are directly wired and interconnected with each other:

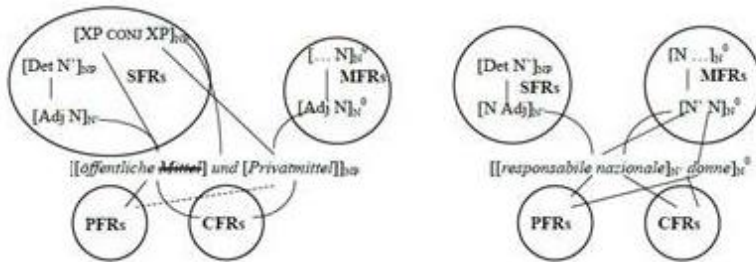


Figure 5. **CAPTION NEEDED HERE**

Moreover, it is important to emphasize the Emergentist approach adopted here which implies that the single modules (and, for our purposes, the Lex₂) do not exist a priori as in the traditional modular view of the grammar which is commonly adopted by most Lexicalist approaches (cf. Scalise and Guevara 2005), but they emerge as generalizations of concrete constructions providing schemas of gradually increasing abstractness, hierarchically connected by means of default inheritance as shortly represented in Figure 5, very much in the sense of Booij's (2010) Construction Morphology. Finally, in the quadripartite architecture sketched in Figure 4, I maintain with Jackendoff (2002, 2013) an independent module of Conceptual formation rules which is made responsible for the semantic aspects associated with the constructions. It remains to be understood whether this is empirically adequate, as suggested by Jackendoff, or whether a more holistic view of conceptualization has to be assumed which makes it superfluous to have all conceptualization cast into a single module. This is left for future research.

References

- Anderson, Stephen. *A-Morphous Morphology*. Cambridge: Cambridge University Press, 1992.
- Apothéloz, Denis. *La construction du lexique français*. Paris: Ophrys, 2002.
- Arnaud, Pierre J. L. *Les composés timbre-poste*. Lyons: Presses Universitaires Lyon, 2003.
- Aronoff, Mark. *Word-Formation in Generative Grammar*. Cambridge, MA: MIT Press, 1976.

- Aronoff, Mark. *Morphology by Itself*. Cambridge, MA: MIT Press, 1994.
- Askedal, John Ole. Interkategoriale Koordination von Adjektiv und Erstelement eines Kompositums im gegenwärtigen Deutsch im Vergleich mit den Norwegischen und dem Englischen. In *Deutsche Syntax: Empirie und Theorie*, edited by Franz J. D'Avis, 1-23. Göteborg: Acta Universitatis Gothoburgensis, 2005.
- Baayen, Harald R. "Corpus linguistics in morphology: Morphological productivity." In *Corpus Linguistics. An International Handbook*, vol. 2, edited by Anke Lüdeling and Merja Kytö, 899-919. Berlin and New York: Mouton de Gruyter, 2009.
- Barlow, Michael and Suzanne Kemmer (Eds.) *Usage Based Models of Grammar*. Stanford, CA: CSLI Publications, 2000.
- Baroni, Marco, Emiliano Guevara and Roberto Zamparelli. "The dual nature of deverbal nominal constructions." *Corpus Linguistics and Linguistic Theory* 5 (1) (2009): 27-60.
- Bauer, Laurie. *English Word-Formation*. Cambridge: Cambridge University Press, 1983.
- Bauer, Laurie. *Morphological Productivity*. Cambridge: Cambridge University Press, 2001.
- Booij, Geert. "Coordination reduction in complex words: A case for prosodic phonology." In *Advances in nonlinear phonology*, edited by Harry van der Hulst and Norval Smith, 143-160. Dordrecht: Foris, 1985.
- Booij, Geert. "Lexical integrity as a formal universal: A constructionist view." In *Universals of Language Universal Today*, edited by Sergio Scalise, Antonietta Bisetto and Elisabetta Magni, 83-100. Dordrecht: Springer, 2009.
- Booij, Geert. *Construction Morphology*. Oxford: Oxford University Press, 2010.
- Corbin, Danielle. "Locutions, composés, unités polylexématiques: lexicalisation et mode de construction." In *La locution entre langue et usages*, edited by Michel Martins-Baltar, 53-101. Fontenay: École Normale Supérieure, 1997.
- Di Sciullo, Anna Maria and Edwin Williams. *On the Definition of Word*. Cambridge, MA: MIT Press, 1987.
- Dressler, Wolfgang Ulrich. "Preferences vs. strict universals in morphology: Word-based rules." In *Theoretical Morphology*, edited by Michael Hammond and Michael Noonan, 143-153. San Diego, CA: Academic Press, 1988.
- Embick, David and Rolf Noyer. "Distributed morphology and the syntax-morphology interface." In *The Oxford Handbook of Linguistic*

- Interfaces*, edited by Gillian Ramchand and Charles Reiss, 290-324. Oxford: Oxford University Press, 2007.
- Fradin, Bernard. "IE, Romance: French." In *The Oxford Handbook of Compounding*, edited by Rochelle Lieber and Pavol Štekauer, 417-435. Oxford: Oxford University Press, 2009.
- Gaeta, Livio. "Ai limiti della morfologia basata sulle parole." In *Scritti di morfologia. In onore di Sergio Scalise in occasione del suo 60° compleanno*, edited by Antonietta Bisetto, Claudio Iacobini and Anna M. Thornton, 47-59. Roma: Caissa Italia, 2003.
- Gaeta, Livio. "Exploring grammaticalization from below." In *What Makes Grammaticalization? A Look from its Fringes and its Components*, edited by Walter Bisang, Nikolaus P. Himmelmann and Björn Wiemer, 45-75. Berlin and New York: Mouton de Gruyter, 2004.
- Gaeta, Livio. "Lexical integrity as a constructional strategy." *Lingue e Linguaggio* 5 (1) (2006): 1-16.
- Gaeta, Livio. "Constituent order in compounds and syntax: Typology and diachrony." *Morphology* 18 (2) (2008): 117-141.
- Gaeta, Livio. "Restrictions in word formation." In *Word-formation. An international handbook of the languages of Europe, vol. 2*, edited by Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen and Franz Rainer, 858-874. Berlin and New York: Mouton de Gruyter, 2015a.
- Gaeta, Livio. "Lexeme formation in a conscious approach to the lexicon." In *Semantics of Complex Words*, edited by Laurie Bauer, Lívia Körtvélyessy and Pavol Štekauer, 115-141. Dordrecht: Springer, 2015b.
- Gaeta, Livio. "Evaluative morphology and sociolinguistic variation." In *Edinburgh Handbook of Evaluative Morphology*, edited by Nicola Grandi and Lívia Körtvélyessy, 121-133. Edinburgh: Edinburgh University Press, 2015c.
- Gaeta, Livio and Davide Ricca. "Productivity in Italian word formation: A variable-corpus approach." *Linguistics* 44 (1) (2006): 57-89.
- Gaeta, Livio and Davide Ricca. "Productivity." In *Word-Formation. An International Handbook of the Languages of Europe, vol. 2*, edited by Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen and Franz Rainer, 841-858. Berlin and New York: Mouton de Gruyter, 2015.
- Geeraerts, Dirk. *Cognitive Semantics*. Oxford: Oxford University Press, 2010.
- Goldberg, Adele E. *Constructions at Work*. Cambridge: Cambridge University Press, 2006.
- Goldberg, Adele E. "Constructionist approaches to language." In *The Oxford Handbook of Construction Grammar*, edited by Thomas

- Hoffmann and Graeme Trousdale, 15-31. Oxford: Oxford University Press, 2013.
- Gries, Stefan Th. "Data in Construction Grammar." In *The Oxford Handbook of Construction Grammar*, edited by Thomas Hoffmann and Graeme Trousdale, 93-108. Oxford: Oxford University Press, 2013.
- Hopper, Paul. "Emergent grammar." In *Berkeley Linguistics Society 13: General Session and Parasession on Grammar and Cognition*, edited by Jon Aske, Natasha Berry, Laura Michaelis and Hana Filip, 139-157. Berkeley, CA: Berkeley Linguistics Society, 1987.
- Jackendoff, Ray. *Foundations of Language*. Oxford: Oxford University Press, 2002.
- Jackendoff, Ray. "Constructions in the parallel architecture." In *The Oxford Handbook of Construction Grammar*, edited by Thomas Hoffmann and Graeme Trousdale, 70-92. Oxford: Oxford University Press, 2013.
- Langacker, Ronald Wayne. *Foundations of Cognitive Grammar, vol. 1*. Stanford, CA: Stanford University Press, 1987.
- Lawrenz, Brigitte. "Der Zwischen-den-Mahlzeiten-Imbiß und der Herren-der-Welt-Größenwahn: Aspekte der Struktur und Bildungsweisen von Phrasenkomposita im Deutschen." *Zeitschrift für Germanistische Linguistik* 24 (1996): 1-15.
- Meibauer, Jörg. "Phrasenkomposita zwischen Wortsyntax und Lexikon." *Zeitschrift für Sprachwissenschaft* 22 (2) (2003): 153-188.
- Meibauer, Jörg. "How marginal are phrasal compounds? Generalized insertion, expressivity and I/Q interaction." *Morphology* 17 (2007): 233-259.
- Mugdan, Joachim. "Units of word-formation." In *Word-Formation. An International Handbook of the Languages of Europe, vol. 1*, edited by Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen and Franz Rainer, 235-301. Berlin and New York: Mouton de Gruyter, 2015.
- Plag, Ingo. *Word-Formation in English*. Cambridge: Cambridge University Press, 2003.
- Scalise, Sergio. *Generative Morphology*. Dordrecht: Foris, 1984.
- Scalise, Sergio. *Morfologia*. Bologna: Il Mulino, 1994.
- Scalise, Sergio and Emiliano Guevara. "The Lexicalist approach to word-formation and the notion of the lexicon." In *Handbook of Word-Formation*, edited by Pavol Štekauer and Rochelle Lieber, 147-187. Dordrecht: Springer, 2005.
- Scalise, Sergio and Antonietta Bisetto. *La struttura delle parole*. Bologna: Il Mulino, 2008.

- Selkirk, Elizabeth O. *The Syntax of Words*. Cambridge, MA: MIT Press, 1982.
- Wiese, Richard. "Phrasal compounds and the theory of word syntax." *Linguistic Inquiry* 27 (1996): 183-193.
- Wischer, Ilse. "Grammaticalization and word formation." In *The Oxford Handbook of Grammaticalization*, edited by Bernd Heine and Heiko Narrog, 356-364. Oxford: Oxford University Press, 2011.